What a smoothie!

Student Worksheet

TI-Nspire CAS   How-to   Student   60 min

Introduction

In this task, you will construct a template for smoothing a time series, and representing it visually. It takes a little while to construct, but if saved it can be used to help with smoothing series investigations, regardless of the size of the data set. The following smoothing methods will be included in the template:

| Mean based	Median based
• *3-point moving mean*	• *3-point moving median*
• *5-point moving mean*	• *5-point moving median*
• *4-point moving mean (centred)*	• *4-point moving mean (centred)*

Construct a spreadsheet template for smoothed variables

The first job is to construct a spreadsheet template that allows the user to enter time series data, and have the smoothed variable values calculated, observable and graphable. To check the success of our template, we will first enter and use the following time series data:

time	1	2	3	4	5	6	7	8
yval	120	150	100	180	200	240	235	245

To enter the data on a spreadsheet and create a time series plot using the TI-Nspire CAS:

- Press ⌂ on > **New Document** and select **Add Lists & Spreadsheet**.
- Press ⌂ on > **Settings > Document Settings** to adjust settings so that **Display Digits = Fix 2** and **Calculation Mode = Approximate**.
- Click **OK** to save the settings changes, and then select **Current** to return to the spreadsheet.
- Locate the cursor at the top of column A and type the heading "**time**"
- Locate the cursor at the top of column B and type the heading "**yval**"
- Enter the data in the first two columns (should appear as shown right)
- Press ctrl doc▾ and select **Add Data & Statistics**.
- Press tab to select **time** as *explanatory* variable
- Press tab to select **yval** as *response* variable
- Press menu > **Plot Type >X-Y Line Plot** (result shown right)

Author: D. Tynan

TEXAS INSTRUMENTS
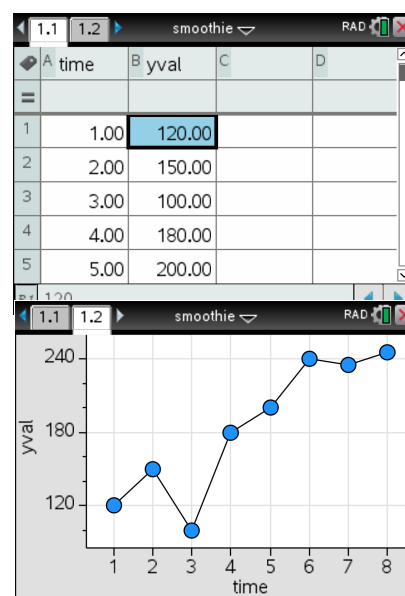
Constructing formulae for 3-point and 5-point moving mean values

It would be useful if the spreadsheet was able to work correctly when a time series with a different number of values is used. To do this, we will use a formula which 'checks' how many values there are in the current series, and displays all relevant smoothed values. For example, consider the formula below, which is to be stored in cell C2

$$=iffn(a2<=dim(time)-1,mean(b1:b3),"")$$

This command can be interpreted in the following manner:

"If the contents of cell A2 is a number less than or equal to 7 (i.e. one less than the number of values in the time series), then display the mean of cells B1, B2 and B3, otherwise display nothing"
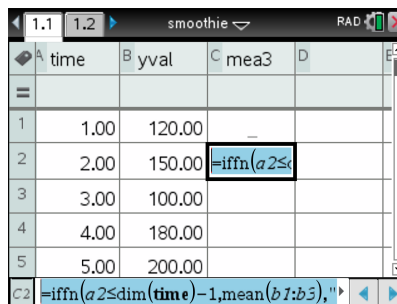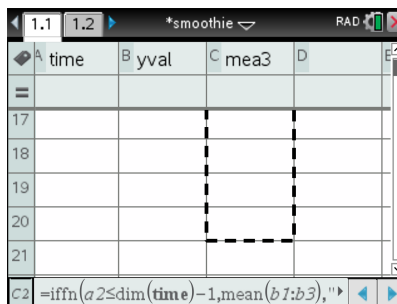
Question 1.

Look carefully at the formula above typed into cell C2. Why is the formula not entered into C1?

The formula entered in cell C2 can be filled down as far as needed so that it finds the mean of the correct 3 nearby values of the time series for that time value. For our purposes, we will fill down so that the template works for a time series with up to 20 time series values.

To do this using the TI-Nspire CAS:

- Locate the cursor at the top of column C and type the heading "**mea3**"

- Locate the cursor in cell C2 and type in the formula
  $=iffn(a2<=dim(time)-1,mean(b1:b3),"")$
  [Note: the "quotes" character can be entered by pressing
  [ctrl] [✗])

- Press [enter] to calculate the value obtained from this formula (the value displayed should be 123.33)

- Locate the cursor again in cell C2

- To fill the formula down in column C, press [menu] > **Data** > **Fill**

- Press ▼ until you reach cell C20, and then press [enter] (the formula will now have been filled down to C20)

Note that the cells from C8 to C20 are blank, as the values in the corresponding cells in column A are not less than or equal to 7.
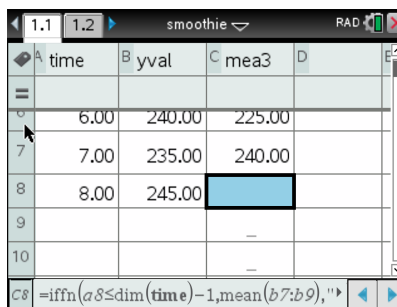
Question 2.

Write a similar formula to be placed in cell D3 that will return the first 5-point moving mean smoothed value. [Hint: remember it should only return a smoothed value from the 3rd value to the 3rd last value in the series].

Author: D. Tynan

TEXAS INSTRUMENTS

Once you are happy that your answer to Question 2 is correct, go to the top of column D and type the column heading "**mea5**", and then enter your formula in cell D3.
Then fill down to cell D20.

The screen should resemble the one shown at right.

	A time	B yval	C mea3	D mea5	E
=					
1	1.00	120.00	_	_	
2	2.00	150.00	123.33	_	
3	3.00	100.00	143.33	=iffn(a3<(	
4	4.00	180.00	160.00	174.00	
5	5.00	200.00	206.67	191.00	

D3 $=iffn(a3<dim(time)-2,mean(b1:b5),"$

Constructing formulae for 3-point and 5-point moving median values
The process is fairly similar for 3-pont and 5-point moving medians. However, we will need to use the inbuilt median command.

In column E, type a column heading "med3" and then construct a formula in E2 that will return the first 3-point moving median smoothed value. Then fill the formula down to cell E20.

Question 3.
What formula did you use in E2? What value is obtained from this formula?

Question 4.
In column F, type a column heading "med5" and then construct a formula in F3 that will return the first 5-point moving median smoothed value. Then fill the formula down to cell F20.

Here is the spreadsheet after the extra formulae have been added. Press ctrl S to save your work.

	A time	B yval	C mea3	D mea5	E med3	F med5
=						
1	1.00	120.00	_	_	_	_
2	2.00	150.00	123.33	_	120.00	_
3	3.00	100.00	143.33	150.00	150.00	150.00
4	4.00	180.00	160.00	174.00	180.00	180.00
5	5.00	200.00	206.67	191.00	200.00	200.00
6	6.00	240.00	225.00	220.00	235.00	235.00
7	7.00	235.00	240.00		240.00	
8	8.00	245.00				
9			_	_	_	_
10			_	_	_	_

Author: D. Tynan

TEXAS INSTRUMENTS

Constructing formulae for 4-point and centred 4-point moving mean values

First consider a 4-point smoothed value, which is inconveniently centred halfway between the 2nd and 3rd time period (i.e. *time* = 2.5). Recall that to deal with this, adjacent 4-point moving mean values can be centred on an actual time period via a further 2-point moving mean smoothing process. For example, this extra smoothing would centre 4-point moving mean values centred on *time* = 2.5 and 3.5 so that it is centred on *time* = 3.

In terms of the spreadsheet, we will construct a formula for the 4-point moving mean values which is not correctly centred (for example, it's harder to introduce an extra line for time = 2.5 and 3.5 value!). However the formula will be correctly aligned again for the centred 4-point moving mean values.

To do this using the TI-Nspire CAS :

- Locate the cursor at the top of column G and type a column heading "**mea4**"

- Locate the cursor in the cell **G2** and type in the formula *=iffn(a2<=dim(time)-2,mean(b1:b4),"")*

- Press ⌊enter⌋ to calculate the value obtained from this formula (the value displayed should be 137.50)

- Fill the formula down to G20

Note that these 4-point moving mean values are **NOT** correctly aligned with the time periods. The values will be correct, but the time period would need to be reduced by 0.5 to align them correctly. This issue is resolved for the centred 4-point moving mean smoothed values in the next step.

To create the centred 4-point smoothed moving mean values using the TI-Nspire CAS :

- Locate the cursor at the top of column H and type a column heading "**cmea4**"

- Locate the cursor in the cell **H3** and type in the formula *=iffn(a3<=dim(time)-2,(g2+g3)/2,"")*

- Press ⌊enter⌋ to calculate the value obtained from this formula (the value displayed should be 147.50)

- Fill the formula down to **H20**

Constructing formulae for 4-point and centred 4-point moving median values
Question 5.

Repeat the above process to calculate both the 4-point moving *median* and centred 4-point moving *median* values.

Author: D. Tynan

TEXAS INSTRUMENTS

Using the "smoothie" template

If all formulas are correct, the following results should be visible for the dataset provided. Check that your template has generated the same smoothed values. Then save the file again (**CTRL-S**)

| A time	B yval	C mea3	D mea5	E med3	F med5	G mea4	H cmea4	I med4	J cmed4
1.00	120.00	_	_	_	_	_	_	_	_
2.00	150.00	123.33	_	120.00	_	137.50	_	135.00	_
3.00	100.00	143.33	150.00	150.00	150.00	157.50	147.50	165.00	150.00
4.00	180.00	160.00	174.00	180.00	180.00	180.00	168.75	190.00	177.50
5.00	200.00	206.67	191.00	200.00	200.00	213.75	196.88	217.50	203.75
6.00	240.00	225.00	220.00	235.00	235.00	230.00	221.88	237.50	227.50
7.00	235.00	240.00		240.00					
8.00	245.00								

The template can be used to:

- find the correct smoothed values based on a given time period, or range of time periods
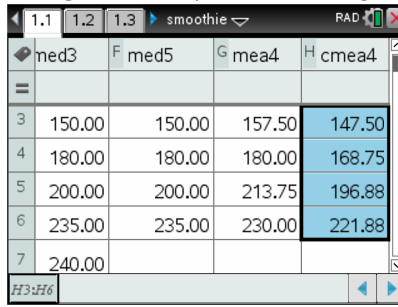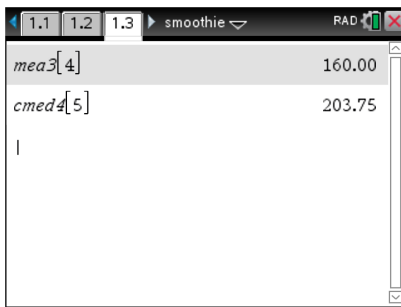
- visually compare the effects and suitability of various smoothing techniques
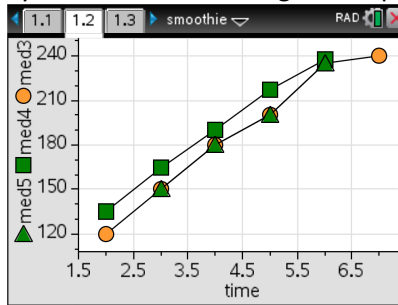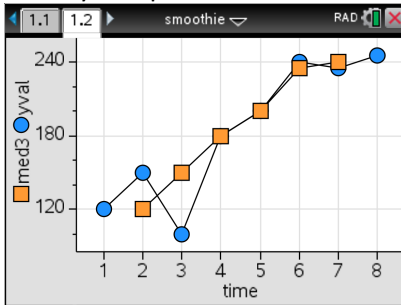
- create smoothed bivariate data sets upon which regression methods may be applied (would need to be copied and pasted to new lists to ensure that the two variables had same number of values, and lined up with time period correctly
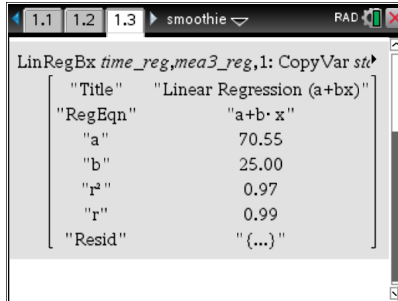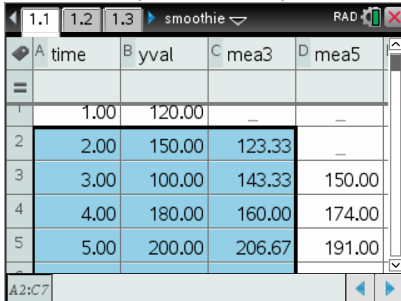
VIP Note!: To use the template with a new time series, select the **time** column and press UP arrow until is is all selected. Then press **MENU-3-4** to clear the data (this does not delete the variable name, which is important as these names are used in the formulas in other columns). Repeat for the **yval** column, then enter the data for the new time series.

Author: D. Tynan

TEXAS INSTRUMENTS